

=====

Sequence Listing was accepted.

If you need help call the Patent Electronic Business Center at (866) 217-9197 (toll free).

Reviewer: Anne Corrigan

Timestamp: [year=2008; month=12; day=10; hr=8; min=36; sec=12; ms=270;]

=====

Application No: 10591029 Version No: 1.0

Input Set:

Output Set:

Started: 2008-12-09 15:33:17.645
Finished: 2008-12-09 15:33:20.383
Elapsed: 0 hr(s) 0 min(s) 2 sec(s) 738 ms
Total Warnings: 49
Total Errors: 0
No. of SeqIDs Defined: 49
Actual SeqID Count: 49

Error code	Error Description
W 213	Artificial or Unknown found in <213> in SEQ ID (1)
W 213	Artificial or Unknown found in <213> in SEQ ID (2)
W 213	Artificial or Unknown found in <213> in SEQ ID (3)
W 213	Artificial or Unknown found in <213> in SEQ ID (4)
W 213	Artificial or Unknown found in <213> in SEQ ID (5)
W 213	Artificial or Unknown found in <213> in SEQ ID (6)
W 213	Artificial or Unknown found in <213> in SEQ ID (7)
W 213	Artificial or Unknown found in <213> in SEQ ID (8)
W 213	Artificial or Unknown found in <213> in SEQ ID (9)
W 213	Artificial or Unknown found in <213> in SEQ ID (10)
W 213	Artificial or Unknown found in <213> in SEQ ID (11)
W 213	Artificial or Unknown found in <213> in SEQ ID (12)
W 213	Artificial or Unknown found in <213> in SEQ ID (13)
W 213	Artificial or Unknown found in <213> in SEQ ID (14)
W 213	Artificial or Unknown found in <213> in SEQ ID (15)
W 213	Artificial or Unknown found in <213> in SEQ ID (16)
W 213	Artificial or Unknown found in <213> in SEQ ID (17)
W 213	Artificial or Unknown found in <213> in SEQ ID (18)
W 213	Artificial or Unknown found in <213> in SEQ ID (19)
W 213	Artificial or Unknown found in <213> in SEQ ID (20)

Input Set:

Output Set:

Started: 2008-12-09 15:33:17.645
Finished: 2008-12-09 15:33:20.383
Elapsed: 0 hr(s) 0 min(s) 2 sec(s) 738 ms
Total Warnings: 49
Total Errors: 0
No. of SeqIDs Defined: 49
Actual SeqID Count: 49

Error code	Error Description
	This error has occurred more than 20 times, will not be displayed

SEQUENCE LISTING

<110> WOOD, DAVID W.
HSII, JUDY
OAK, SEACHOL
CONTRERAS, LYDIA
CHESTNUT, JOHN

<120> SELF-CLEAVING AFFINITY TAGS AND METHODS OF USE

<130> 331772-00103

<140> 10591029
<141> 2008-12-09

<150> PCT/US05/05763
<151> 2005-02-24

<150> 60/548,092
<151> 2004-02-27

<160> 49

<170> PatentIn version 3.3

<210> 1
<211> 504
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
mini-intein derived from the full-length Mycobacterium
tuberculosis recA intein sequence

<400> 1
gccctcgca agggcactcg gatcttcgat ccggtcaccc gtacaacgc tcgcattcgag 60
gatgttgtcg gtggcgcaa gcctattcat gtcgtggctg ctgccaaggga cggAACGCTG 120
catgcgcggc ccgtgggtgc ctggttcgac caggaaacgc gggatgtgat cgggttgcgg 180
atcgccggtg gcccattcct gtggcgaca cccgatcaca aggtgctgac agagtacggc 240
tggcgtgccc ccgggaaact ccgcaaggga gacagggtgg cgcaaccgc acgcttcgat 300
ggattcggtg acagtgcgc gattccggcg cgctgcagg cgctcgccg tgccctggat 360
gacaaattcc tgcacgacat gctggcgaa gaactccgat attccgtgat ccgagaagtg 420
ctgccaacgc ggcgggcacg aacgttcggc ctgcgggtcg aggaactgca caccctcgcc 480
gccgaagggg ttgttgtaca caac 504

<210> 2

<211> 168
<212> PRT
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
mini-intein derived from the full-length Mycobacterium
tuberculosis recA intein sequence

<400> 2

Ala Leu Ala Glu Gly Thr Arg Ile Phe Asp Pro Val Thr Gly Thr Thr
1 5 10 15

His Arg Ile Glu Asp Val Val Gly Gly Arg Lys Pro Ile His Val Val
20 25 30

Ala Ala Ala Lys Asp Gly Thr Leu His Ala Arg Pro Val Val Ser Trp
35 40 45

Phe Asp Gln Gly Thr Arg Asp Val Ile Gly Leu Arg Ile Ala Gly Gly
50 55 60

Ala Ile Leu Trp Ala Thr Pro Asp His Lys Val Leu Thr Glu Tyr Gly
65 70 75 80

Trp Arg Ala Ala Gly Glu Leu Arg Lys Gly Asp Arg Val Ala Gln Pro
85 90 95

Arg Arg Phe Asp Gly Phe Gly Asp Ser Ala Pro Ile Pro Ala Arg Val
100 105 110

Gln Ala Leu Ala Asp Ala Leu Asp Asp Lys Phe Leu His Asp Met Leu
115 120 125

Ala Glu Glu Leu Arg Tyr Ser Val Ile Arg Glu Val Leu Pro Thr Arg
130 135 140

Arg Ala Arg Thr Phe Gly Leu Glu Val Glu Leu His Thr Leu Val
145 150 155 160

Ala Glu Gly Val Val Val His Asn
165

<210> 3

<211> 504

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
TOPO modified intein sequence from SEQ ID NO: 1

<400> 3

gcccctcgca ggggcactcg gatcttcgat ccgggtcaccg gtacaacgca tcgcacatcgag 60
gatgttgtcg gtgggcgcaa gcctattcat gtcgtggctg ctgccaaggga cggaaacgctg 120
catgcgcggc ccgtgggtgc ctgggttcgac caggaaacgc gggatgtgat cgggttgcgg 180
atcgccggtg gcgcacatcct gtgggcgaca cccgatcaca aggtgctgac agagtacggc 240
tggcgtgccc ccgggaaact ccgcaaggga gacagggtgg cgcaaccgca acgcttcgat 300
ggatccggtg acagtgcgcc gatccggcg cgcgtgcagg cgctcgccga tgccctggat 360
gacaaattcc tgcacgacat gctggcggaa gaactccgct attccgtgat ccgagaagtg 420
ctgccaacgc ggccggcact aacgttcggc ctgcgggtcg aggaactgca caccctcgac 480
gccgaagggg tccttgtaca caac 504

<210> 4

<211> 168

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
TOPO modified intein sequence from SEQ ID NO: 2

<400> 4

Ala Leu Ala Glu Gly Thr Arg Ile Phe Asp Pro Val Thr Gly Thr Thr
1 5 10 15

His Arg Ile Glu Asp Val Val Gly Gly Arg Lys Pro Ile His Val Val
20 25 30

Ala Ala Ala Lys Asp Gly Thr Leu His Ala Arg Pro Val Val Ser Trp
35 40 45

Phe Asp Gln Gly Thr Arg Asp Val Ile Gly Leu Arg Ile Ala Gly Gly
50 55 60

Ala Ile Leu Trp Ala Thr Pro Asp His Lys Val Leu Thr Glu Tyr Gly
65 70 75 80

Trp Arg Ala Ala Gly Glu Leu Arg Lys Gly Asp Arg Val Ala Gln Pro

85

90

95

Arg Arg Phe Asp Gly Phe Gly Asp Ser Ala Pro Ile Pro Ala Arg Val
100 105 110

Gln Ala Leu Ala Asp Ala Leu Asp Asp Lys Phe Leu His Asp Met Leu
115 120 125

Ala Glu Glu Leu Arg Tyr Ser Val Ile Arg Glu Val Leu Pro Thr Arg
130 135 140

Arg Ala Arg Thr Phe Gly Leu Glu Val Glu Glu Leu His Thr Leu Val
145 150 155 160

Ala Glu Gly Val Leu Val His Asn
165

<210> 5

<211> 531

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
GATEWAY modified intein sequence from SEQ ID NO: 1

<400> 5

gccctcgca agggcactcg gatcttcgat ccggtcaccc gtacaacgca tcgcacgag 60

gatgttgcgtc gtggggcaca gcctattcat gtctggctg ctgccaaggga cggAACGCTG 120

catgcgcggc ccgtgggtgc ctggttcgac cagggAACGC gggatgtat cgggttgcgg 180

atcgccggtg gcccattct gtggggcaca cccgatcaca aggtgctgac agagtacggc 240

tggcgtgccc ccgggaaact ccgcaggaga gacagggtgg cgcaaccgac acgtttcgat 300

ggattcggtg acagtgcgcc gattccgaca agtttgtaca aaaaaggcagg cagcgcgcgc 360

gtgcaggcgc tcgcggatgc cctggatgac aaattcctgc acgacatgct ggccggaa 420

ctccgctatt ccgtgatccg agaagtgcgt ccaacgcggc gggcacgaaac gttcggcctc 480

gaggtcgagg aactgcacac cctcgtcgac gaagggttg ttgtacacaa c 531

<210> 6

<211> 177

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
GATEWAY modified intein sequence from SEQ ID NO: 2

<400> 6

Ala Leu Ala Glu Gly Thr Arg Ile Phe Asp Pro Val Thr Gly Thr Thr
1 5 10 15

His Arg Ile Glu Asp Val Val Gly Gly Arg Lys Pro Ile His Val Val
20 25 30

Ala Ala Ala Lys Asp Gly Thr Leu His Ala Arg Pro Val Val Ser Trp
35 40 45

Phe Asp Gln Gly Thr Arg Asp Val Ile Gly Leu Arg Ile Ala Gly Gly
50 55 60

Ala Ile Leu Trp Ala Thr Pro Asp His Lys Val Leu Thr Glu Tyr Gly
65 70 75 80

Trp Arg Ala Ala Gly Glu Leu Arg Lys Gly Asp Arg Val Ala Gln Pro
85 90 95

Arg Arg Phe Asp Gly Phe Gly Asp Ser Ala Pro Ile Pro Thr Ser Leu
100 105 110

Tyr Lys Lys Ala Gly Ser Ala Arg Val Gln Ala Leu Ala Asp Ala Leu
115 120 125

Asp Asp Lys Phe Leu His Asp Met Leu Ala Glu Glu Leu Arg Tyr Ser
130 135 140

Val Ile Arg Glu Val Leu Pro Thr Arg Arg Ala Arg Thr Phe Gly Leu
145 150 155 160

Glu Val Glu Glu Leu His Thr Leu Val Ala Glu Gly Val Val Val His
165 170 175

Asn

<210> 7

<211> 525

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
TOPO and GATEWAY modified intein sequence from
SEQ ID NO: 1

<400> 7

gccctcgca agggcactcg gatcttcgat ccggtcaccc gtacaacgca tcgcacatcgag 60
gatgttgcg gtggcgcaa gcctattcat gtcgtggctg ctgccaagga cggAACGCTG 120
catgcgcggc ccgtgggtgc ctgggttcgac caggaaacgc gggatgtgat cgggttgcgg 180
atcgccggtg gcccattcct gtgggcgaca cccgatcaca aggtgctgac agagtacggc 240
tggcgtgccc ccgggaaact ccgcaaggga gacagggtgg cgcaaccgac acgcttcgat 300
ggattcggtg acagtgcgac gattccgaca agtttgtaca aaaaaggcagg cagcgcgcgc 360
gtgcaggcgc tcgcggatgc cctggatgac aaattcctgc acgacatgct ggccgaagaa 420
ctccgctatt ccgtgatccg agaagtgtcg ccaacgcggc gggcacgaac gttcggcctc 480
gaggaactgc acacctcgt cggccgaaggg gtccttgtac acaac 525

<210> 8

<211> 177

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
TOPO and GATEWAY modified intein sequence from
SEQ ID NO: 2

<400> 8

Ala Leu Ala Glu Gly Thr Arg Ile Phe Asp Pro Val Thr Gly Thr Thr
1 5 10 15

His Arg Ile Glu Asp Val Val Gly Gly Arg Lys Pro Ile His Val Val
20 25 30

Ala Ala Ala Lys Asp Gly Thr Leu His Ala Arg Pro Val Val Ser Trp
35 40 45

Phe Asp Gln Gly Thr Arg Asp Val Ile Gly Leu Arg Ile Ala Gly Gly
50 55 60

Ala Ile Leu Trp Ala Thr Pro Asp His Lys Val Leu Thr Glu Tyr Gly
65 70 75 80

Trp Arg Ala Ala Gly Glu Leu Arg Lys Gly Asp Arg Val Ala Gln Pro
85 90 95

Arg Arg Phe Asp Gly Phe Gly Asp Ser Ala Pro Ile Pro Thr Ser Leu
100 105 110

Tyr Lys Lys Ala Gly Ser Ala Arg Val Gln Ala Leu Ala Asp Ala Leu
115 120 125

Asp Asp Lys Phe Leu His Asp Met Leu Ala Glu Glu Leu Arg Tyr Ser
130 135 140

Val Ile Arg Glu Val Leu Pro Thr Arg Arg Ala Arg Thr Phe Gly Leu
145 150 155 160

Glu Val Glu Glu Leu His Thr Leu Val Ala Glu Gly Val Leu Val His
165 170 175

Asn

<210> 9
<211> 6262
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
pET-GWMIT sequence

<220>
<221> modified_base
<222> (817)..(820)
<223> a, c, g, t, unknown or other

<220>
<221> modified_base
<222> (822)..(822)
<223> Inosine

<220>
<221> modified_base
<222> (823)..(826)
<223> a, c, g, t, unknown or other

<220>
<221> modified_base
<222> (827)..(828)
<223> Inosine

<220>
<221> modified_base
<222> (829)..(841)
<223> a, c, g, t, unknown or other

<400> 9

caaggagatg	gcgcccaaca	gtccccggc	cacggggct	gccaccatac	ccacgcccga	60
acaagcgctc	atgagccga	agtggcgagc	ccgatcttcc	ccatcggtga	tgtcggcgat	120
ataggcgcca	gcaaccgcac	ctgtggcgcc	ggtgatgccg	gccacgatgc	gtccggcgta	180
gaggatcgag	atctcgatcc	cgcgaaattta	atacgactca	ctatagggga	atttgtgagcg	240
gataacaatt	cccctctaga	aataattttg	tttaacttta	agaaggaatt	cgccctcgca	300
gagggcactc	ggatcttcga	tccggtcacc	ggtacaacgc	atcgcatcga	ggatgttgc	360
ggtgggcgca	agcctattca	tgtcggtggct	gctgccaagg	acggaacgct	gcatgcgcgg	420
cccggtggtgt	cctggttcga	ccagggaaacg	cgggatgtga	tccggttgcg	gatcgccggt	480
ggcgccatcc	tgtggcgac	acccgatcac	aaggtgctga	cagagtacgg	ctggcgtgcc	540
gccggggaaac	tccgcaaggg	agacagggtg	gchgcaaccgc	gacgcttcga	tggattcggt	600
gacagtgcgc	cgattccgac	aagtttgta	aaaaaagcag	gcagcgcgcg	cgtgcaggcg	660
ctcgccggatg	ccctggatga	caaattctg	cacgacatgc	tggcggaaaga	actccgctat	720
tccgtgatcc	gagaagtgt	gccaacgcgg	cgggcacgaa	cgttcggcct	cgaggaactg	780
cacaccctcg	tcgccaagg	ggttcttgt	cacaacnnnn	mnnnnnnnnnnn	nnnnnnnnnnnn	840
nggtaagcct	atccctaacc	ctctcctcg	tctcgattct	acgcgtaccg	gtcatcatca	900
ccatcaccat	tgagtttgat	ccggctgcta	acaaagcccg	aaaggaagct	gagttggctg	960
ctgccaccgc	tgagcaataa	ctagcataac	cccttggggc	ctctaaacgg	gtcttgaggg	1020
gtttttgt	gaaaggagga	actatatccg	gatatcccgc	aagaggcccg	gcagtaccgg	1080
cataaccaag	cctatgccta	cagcatccag	ggtgacgggt	ccgaggatga	cgatgagcgc	1140
attgttagat	ttcatacacy	gtgcctgact	gcgttagcaa	tttaactgtg	ataaactacc	1200
gcattaaagc	ttatcgatga	taagctgtca	aacatgagaa	ttaattcttgc	aagacgaaag	1260
ggcctcgta	tacgcctatt	tttataggtt	aatgtcatga	taataatggt	ttcttagacg	1320
tcaggtggca	ctttcgggg	aaatgtgcgc	ggaacccta	tttgcgttatt	tttctaaata	1380
cattcaaata	tgtatccgct	catgagacaa	taaccctgat	aaatgcttca	ataatattga	1440
aaaaggaaga	gtatgagttat	tcaacatttc	cgtgtcgccc	ttattccctt	ttttgcggca	1500

ttttgccttc ctgttttgc tcacccagaa acgctggta aagtaaaaga tgctgaagat 1560
cagttgggtg cacgagtggg ttacatcgaa ctggatctca acagcggtaa gatccttgag 1620
agtttgcgc ccgaagaacg tttccaatg atgagcactt taaaagttct gctatgtggc 1680
gcggtattat cccgtgtga cgccggcaa gagcaactcg gtcgcccat acactattct 1740
cagaatgact tggtagtgc ctcaccaggc acagaaaaagc atcttacgga tggcatgaca 1800
gtaagagaat tatgcagtgc tgccataacc atgagtgata acactgcggc caacttactt 1860
ctgacaacga tcggaggacc gaaggagcta accgctttt tgccacaacat gggggatcat 1920
gtaactcgcc ttgatcggtt ggaaccggag ctgaatgaag ccataccaaa cgacgagcgt 1980
gacaccacga tgcctgcagc aatggcaaca acgttgcgca aactattaac tggcgaacta 2040
cttactctag cttccggca acaattaata gactggatgg aggcggataa agttgcagga 2100
cacttctgc gctcgccct tccggctggc tggtttattt ctgataaattc tggagccggt 2160
gagcgtgggt ctgcggat catgcagca ctggggccag atggtaagcc ctcccgatc 2220
gtagttatct acacgacggg gagtcaggca actatggatg aacgaaatag acagatcgct 2280
gagataggtg ctcactgtat taagcattgg taactgtcag accaagtta ctcataatata 2340
cttagattt attaaaact tcattttaa tttaaaagga tctaggtgaa gatcctttt 2400
gataatctca tgacaaaaat cccttaacgt gagtttctgt tccactgagc gtcagacccc 2460
gtagaaaaga tcaaaggatc ttcttgagat ctttttttc tgccgtaat ctgctgcttg 2520
caaacaaaaa aaccacccgct accagcggtg gtttgttgc cggatcaaga gctaccaact 2580
cttttccga aggttaactgg cttcagcaga ggcagatac caaatactgt cttcttagtg 2640
tagccgtagt taggccccca cttcaagaac tctgttagcac cgcctacata cctcgctctg 2700
ctaattctgt taccagtggc tgctgcccagt ggcgataagt cgtgtcttac cgggttggac 2760
tcaagacgt agttaccgga taaggcgca gggcggtt gaaacgggggg ttcgtgcaca 2820
cagccccagct tggaggcaac gacctacacc gaactgagat acctacagcg tgagctatga 2880
gaaagcgcca cgcttcccgaa agggagaaag gggacaggt atccggtaag cggcagggtc 2940
ggaacaggag agcgcacgag ggagcttcca gggggaaacg cctggtatct ttatagtcct 3000
gtcggtttc gccacctctg acttgagcgt cgattttgt gatgctcgac agggggggcgg 3060
agcctatgga aaaacgcccag caacgcggcc ttttacggt tcctggcctt ttgctggcct 3120
tttgctcaca tgttcttcc tgctgtatcc cctgattctg tggataaccg tattaccgcc 3180
tttgagtgag ctgataccgc tgcggcagc cgaacgaccg agcgcagcga gtcagtgagc 3240

gaggaagcg aagagcgct gatgcggtat tttctcctta cgcatctgtg cggatttca 3300
caccgcaatg gtgcactctc agtacaatct gctctgatgc cgcatagtta agccagtata 3360
cactccgcta tcgctacgtg actgggtcat ggctgcgccc cgacacccgc caacacccgc 3420
tgacgcgccc tgacgggctt gtctgtccc ggcattccgct tacagacaag ctgtgaccgt 3480
ctccgggagc tgcatgtgtc agaggttttc accgtcatca ccgaaacgcg cgaggcagct 3540
gcggtaaagc tcatcagcgt ggtcgtgaag cgattcacag atgtctgcct gttcatccgc 3600
gtccagctcg ttgagttct ccagaagcgt taatgtctgg cttctgataa agcgggcat 3660
gttaagggcg gtttttctt gtttgtcac tgatgcctcc gtgttaagggg gatttctgtt 3720
catggggta atgataccga taaaacgaga gaggatgctc acgatacggg ttactgatga 3780
tgaacatgcc cggttactgg aacgttgtga gggtaaacaa ctggcggtat ggatgcggcg 3840
ggaccagaga aaaatcactc agggtaatgc ccagcgcttc gttaatacag atgttaggtgt 3900
tccacagggt agccagcagc atcctgcgtat gcagatccgg aacataatgg tgcaggcgc 3960
tgacttccgc gttccagac tttacgaaac acggaaaccg aagaccattc atgttgttgc 4020
tcaggtcgca gacgtttgc agcagcagtc gttcacgtt cgctcgctta tcggtgattc 4080
attctgctaa ccagtaaggc aaccccgcca gcctagccgg gtctcaacg acaggagcac 4140
gatcatgcgc acccggtggcc aggacccaac gctgccccag atgcgcgcg tgcggctgt 4200
ggagatggcg gacgcgtgg atatgttctg ccaagggttg gttgcgtat tcacagttct 4260
ccgcaagaat tgattggctc caattctgg agtggtaat ccgttagcga ggtgccgccc 4320
gctccattc aggtcgaggt ggcccgctc catgcaccgc gacgcaacgc ggggaggcag 4380
acaaggata gggcgccgccc tacaatccat gccaaccgt tccatgtgtc cgccgaggcg 4440
gcataaatcg ccgtgacgt cagcggtcca atgatcgaag ttaggctgtt aagagccgcg 4500
agcgatcctt gaagctgtcc ctgatggtcg tcatctacct gcctggacag catggctgc 4560
aacgcgggca tcccgatgcc gccggaagcg agaagaatca taatggggaa ggccatccag 4620
cctcgctcg cgaacgcccag caagacgtag cccagcgcgt cggccgcat gcccggata 4680
atggcctgtc tctcgccgaa acgtttggtg gcgggaccag tgacgaaggc ttgagcggagg 4740
gcgtgcaaga ttccgaatac cgcaagcgac aggccgatca tcgtcgctt ccagcgaaag 4800
cggtcctcgc cgaaaatgac ccagagcgct gccggcacct gtctacgag ttgcatgata 4860
aagaagacag tcataagtgc ggcgacgata gtcgtcccc gcggccaccg gaaggagctg 4920

actgggttga aggctctcaa gggcatcggt cgagatcccg gtgcctaatg agtgagctaa 4980
cttacattaa ttgcgttgcg ctcactgccc gctttccagt cgggaaacct gtcgtgccag 5040
ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg gcgccagggt 5100
ggttttctt ttcaccagtg agacgggcaa cagctgattg cccttcacccg cctggccctg 5160
agagagttgc agcaagcggt ccacgctggc ttgccccagc ag